

## Machine Learning in 3D Space Gesture Recognition

Veronica Naosekpan\* & Rupam Kumar Sharma

Department of CSE and IT, School of Technology, Assam Don Bosco University, Assam, India

\*Corresponding author: venaosekpan11@gmail.com

Received 24 October 2018, Received in revised form 20 April 2019

Accepted 17 July 2019, Available online 31 October 2019

### ABSTRACT

*The rapid increase in the development of robotic systems in a controlled and uncontrolled environment leads to the development of a more natural interaction system. One such interaction is gesture recognition. The proposed paper is a simple approach towards gesture recognition technology where the hand movement in a 3-dimensional space is utilized to write the English alphabets and get the corresponding output in the screen or a display device. In order to perform the experiment, an MPU-6050 accelerometer, a microcontroller and a Bluetooth for wireless connection are used as the hardware components of the system. For each of the letters of the alphabets, the data instances are recorded in its raw form. 20 instances for each letter are recorded and it is then standardized using interpolation. The standardized data is fed as inputs to an SVM (Support Vector Machine) classifier to create a model. The created model is used for classification of future data instances at real time. Our method achieves a correct classification accuracy of 98.94% for the English alphabets' hand gesture recognition. The primary objective of our approach is the development of a low-cost, low power and easily trained supervised gesture recognition system which identifies hand gesture movement efficiently and accurately. The experimental result obtained is based on use of a single subject.*

**Keywords:** Human Computer Interaction; Gesture Recognition; Machine Learning; Support Vector Machine; Accelerometer.

### INTRODUCTION

The human-computer interaction (HCI) techniques have become so prominent that it has become an indispensable component in our daily life. The primary objective of this technique is to enhance the automated process of machines with very less or limited human intervention. One such technology is the gesture. It is a form of non-verbal means of communication. Gesture recognition is becoming one of the active fields of research as it serves as an intelligent and a natural interface between the human and the computer. With the wider incorporation of accelerometer on consumer electronic devices it has open more scope of device interaction with human gesture.

Gesture recognition can be achieved broadly in two forms – 1) Accelerometer-based approach (Ahmad et al. 2010; Xie et al. 2015; Liu et al. 2009 and 2) Computer Vision-based (Molchanov et al. 2015; Boyali & Kavakli 2012) approach. Conventional computer vision-based hand gesture recognition can track and recognize gestures effectively without any contact with the user. However, vision-based techniques may be affected by lighting conditions of the environment, which in turn limits the application scenarios, and particularly in the mobile environment. Moreover, a camera is needed for this approach which may sometimes be cost crunching. Therefore, going for the accelerometer sensor-based approach is considered as a better alternative in terms cost effectiveness, mobility etc.

A majority of the literatures available on gesture recognition involves accelerometer based since it is cost efficient, has low latency and the underlying architecture of accelerometer is easy to understand. Ahmad et al. (2010) proposed an accelerometer-based gesture recognition system where the gestures are hand movements. It created a dictionary of 18 gestures for a database of about 35000 repetitions collected from 7 users. The core of the training phase is the DTW (Dynamic Time Warping) and the AP (Affinity Propagation) algorithms. Later, the CS (Compressive Sampling) comes into picture. The system has an accuracy of about 95%. It has been found that the one-nearest neighbour DTW is not always efficient and so, the recognition problem is formulated as an  $l_1$ -minimization problem after projecting all candidate gesture traces into a lower dimensional subspace (Ahmad et al. 2011). Xu et al. (2012) uses MEMS (Micro-Electro Mechanical System) accelerometer without gyroscope as heavy computation will be necessary if used with gyroscope. Three gesture recognition models are presented – 1) sign sequence and Hopfield based method, 2) Velocity increment based and 3) Sign sequence and template matching. In all the three mentioned methods, the patterns of the accelerations are segmented directly and recognized in the time domain. Using the simple feature extraction based on sign sequence of acceleration, the system achieves high accuracy without the employment of statistical method such as HMM (Hidden Markov Model). Xie et al. (2015) proposed a similarity matching-based

extensible hand gesture recognition integrated with a smart ring device whose applications include intelligent wheelchair, robot-assisted living and automatic user recognition for television control system. The gesture in this system consists of two types – a basic gesture and a complex gesture. The complex gesture is made up of sequence of basic gestures. The procedure consists of acquisition of acceleration data, pre-processing of the data, gesture segmentation, feature extraction, basic gesture recognition, basic gesture encoding and similarity matching using the template data for complex gesture recognition. In Xie and Cao (2016), the hand gesture recognition is done using the neural network and similarity matching. A pen-type sensing device with feed-forward neural network and similarity matching is presented. The data collected from the acceleration of the hand movement are sent to the personal computer (PC) via the USB cable. The process includes data acquisition and pre-processing, followed by gesture segmentation, feature extraction, classifier construction, basic gesture encoding and lastly, the similarity matching. The segmentation scheme is introduced to identify the starting and end points of the basic gesture automatically. The FNN (Feed-forward Neural Network) classifier provides a good recognition accuracy and the SM (Similarity Matching) approach provides the extendibility features. Also, the system conducts the accuracy evaluation in user-dependent and user-independent environment. Muley and Yadav (2014) proposed a trajectory recognition algorithm on 3-D hand gestures. The steps composed of data acquisition, signal preprocessing, feature generation, feature selection and lastly feature extraction. The algorithm is implemented using a digital pen to construct the gesture pattern. The hand motions are wirelessly transmitted to the computer. It uses the PNN (Probabilistic Neural Network) classifier for dimension reduction of the features. The digital pen based on trajectory recognition algorithm consists of triaxial accelerometer with a 10-b A/D converter and wireless transceiver. uWave (Liu et al. 2009) is an algorithm for HCI (Human Computer Interaction) that uses one set of training sample for each gesture and allows user employ personalized gestures. It uses the accelerometer present in the consumer electronic devices. It matches the unknown gesture readings with the templates using DTW. It is efficient for a resource constrained environment. It achieves an accuracy of 98.6% and 93.5% with and without template adaptation using the gesture vocabulary identified by Nokia research which consists of 4480 gestures collected from 8 participants in multiple weeks. The strength of the uWave is its user-dependent mode which makes it customizable. The driving force for uWave is that human gestures can be represented by the time series of forces applied to the handheld device. The sensing system of Xie et al. (2015) consists of a smart ring which has a 3-axis MEMS accelerometer, a rechargeable battery, Broadcom System on Chip Smart Bluetooth, a free scale semiconductor, a vibrator and LEDs. The components are all mounted on the ring. The Bluetooth reads the accelerometer data through the I2C interface and runs the algorithm simultaneously. The ring is connected to the smartphone via the bluetooth. In Nagadeepa et al. (2016), a trajectory recognition algorithm

has been used for the analysis of inertial sensor data. Gupta et al. (2016) developed an automatic hand gesture spotting algorithm that detects the start and end of a meaningful hand gesture movements. The HCI (Human Computer Interaction) is performed using accelerometer and gyroscope sensors. These hand gestures are used to control smart devices. The recognition of gesture is performed by comparing the gesture template and the gesture database using dynamic time warping algorithm. Six types of gestures are recognized by the system developed. Multiple accelerometers (2 and 5) are used for gesture recognition in Ajay et al. (2017) which are mounted on the fingertips of the user. The system can be configured for N number of users. It has been tested on the American Sign Language (ASL) alphabets. Hsu et al. (2015) introduced a gesture recognition using an inertial sensor-based digital pen and its associated dynamic time warping based algorithm. The highlights of this system is that it reduces the integral error caused by the intrinsic noise of the accelerometer, gyroscope and magnetometer by collecting the signals into a quaternion-based complementary filter. In Setia et al. (2015), a robot is controlled by hand gestures instead of using buttons. Here, there are two sections. In the first section, the transmitting section, the accelerometer is mounted on the hand of the user to capture the user's hand movement in order to drive the robot in motion. The data are encoded using Encoder IC. In the second section, the receiving section, data are decoded using Decoder IC which is again processed by a microprocessor before transmitting the data as a command to the motor driver to rotate the motor of the robot. The primary concern is to move the robot as soon the user makes any recognizable gestures. Compared with the other input devices, accelerometer is easier to work and also, it offers the possibilities of wireless transmission. The other advantages are the small set up time and low cost. In Niezen and Hancke (2009), the existing gesture recognition algorithms are studied and the most appropriate one is selected for further optimization in order to make its work for mobile devices. Jakub et al. (2016) developed a sign language gesture recognition glove based on HMM (Hidden Markov Model) and Parallel HMM approaches. Use of parallel HMM reduced the error rate by more than 60%. It has been found from that most of the existing approaches for gesture recognition are based on HMM (Hidden Markov Model), DTW (Dynamic Time Warping) algorithm, neural networks and uses multiple sensors. Table 1 shows the summary of various existing approaches in gesture recognition techniques.

This paper proposes a gesture recognition system developed using a low-cost MPU-6050 accelerometer and trained using a supervised classification algorithm (Support Vector Machine). The advantage of the proposed model is low power requirement, and a simple circuitry since it does not require special hardware. The proposed application finds application in teaching purposes where the instructor can comfortably sit anywhere in the class room and write the words that the user wants to write in a 3-dimensional space. The classification accuracy obtained is 98.94% which is very competitive with the state-of-the-art approaches

TABLE 1. Tabular comparison of different techniques and their accuracy

No.	Reference	Environment	Accuracy
1	Ahmad Akl et al. (2010)	User-dependent User-independent	99.79% 96.89%
2	Ahmad Akl et al. (2011)	User-dependent User-independent	98.71% 96.84%
3	Ruize et al. (2012)	Non-specific users	95.60%
4	Renqiang Xie et al. (2015)	Basic gestures Complex gestures	98.90% 97.20%
5	Jiayang et al. (2009)	User-dependent (with template) User-dependent (without template)	98.60% 93.50%
6	Yu-Liang et al. (2015)	User-dependent	94.3%
7	Ajay Kannan et al. (2017)	5 accelerometers 2 accelerometers	95.30% 87.00%
8	Hari et al. (2016)	Multi-users	94%
9	Jakub et al. (2016)	User-independent	99.75%

The data collection in this study includes the primary and secondary data. The primary data is obtained by observation, pilot and survey studies while the secondary data is obtained from the college's management regarding to the number of student that registered their vehicles.

## METHODOLOGY

### EXPERIMENT EQUIPMENT

The block diagram of the hardware components of the system is shown in Figure 1.

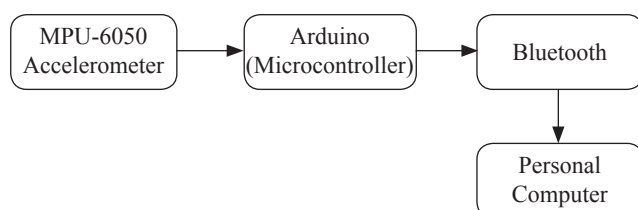


FIGURE 1. Block diagram of the hardware components

**MPU-6050 Accelerometer:** MPU stands for Motion Processing Unit. The MPU-6050 is an IMU (Inertial Measurement Unit) sensors which are compatible with Arduino microcontroller.

It is a 6-DOF (Degree of Freedom) sensor that gives 6 values as its output. It is based on MEMS (Micro Electro Mechanical Systems) technology. The first three values are the output for the 3 axes of the accelerometer and the last 3 values are for the 3 axes of the gyroscope. Both the devices are embedded on a single chip. The working principle for the accelerometer is the Piezoelectric effect and the working principle for the gyroscope Coriolis acceleration. It provides the integrated 6-axis motion processor solution which eliminates the gyroscope and accelerometer cross-axis misalignment which is a common association. It follows the I2C protocol which is a master-slave model where the

accelerometer acts as a slave in most of the experiments. The required input voltage is 2.3-3.4 Volts.

**HC-05 Bluetooth:** It is a popular module for wireless communication. The HC-05 Bluetooth Module can be used as a Master or Slave configuration using I2C communication protocol. Default configuration module active is Slave. The Bluetooth module is a serial port that is fully qualified Bluetooth V2.0 + EDR (Enhanced Data Rate) 3Mbps Modulation with complete 2.4 GHz radio transceiver and baseband. The slave modules can accept connections but cannot initiate a connection to another Bluetooth device in contrary to Master module that can initiates connections to other devices. It requires power supply in the range 3.3V to 6V.

**Arduino UNO:** The Arduino Uno is an open sourced microcontroller board based on the ATmega328 (datasheet). The microcontroller has a 14-digital input/output pins, 6 analog inputs, a 16 MHz ceramic resonator, a USB connection, a power jack, an ICSP header, and a reset button. It is programmable with the Arduino IDE (Integrated Development Environment). The microcontroller present is ATmega328 operating Voltage at 5V and recommended Input Voltage is 7-12V.

**Tactile Switch:** It is an on/off electronic switch that is only on when the button if there is some change in pressure on the button surface. As soon as a tactile switch button is released, the circuit is broken.

Figure 2 shows the circuitry connections of the entire system.

### EXPERIMENT DESIGN AND PROCEDURE

The proposed pipeline of the gesture recognition system is shown in Figure 3.

Different steps of the proposed implementation are as follows:

1. **Data Acquisition:** In this step, the data chunk for each of the letter of the English alphabet are captured separately

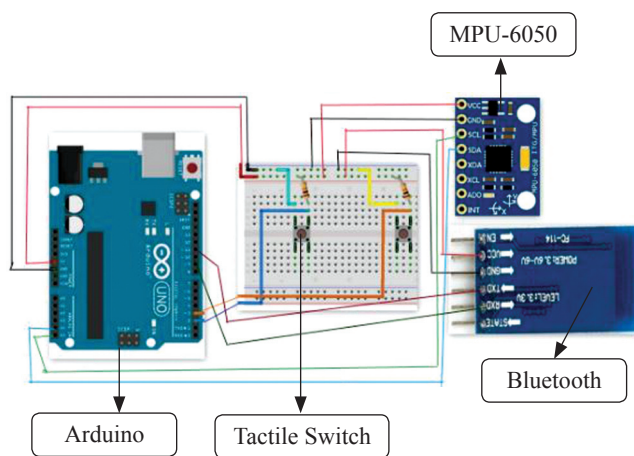


FIGURE 2. Connections of the system

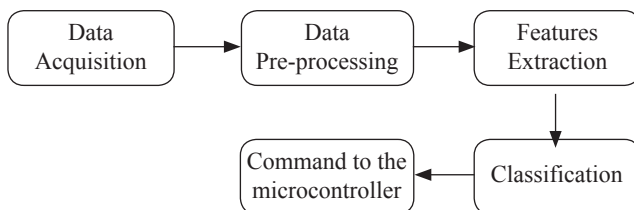


FIGURE 3. Proposed pipeline for hand gesture recognition

in different files. The process of data capturing is done by moving the system with the accelerometer mounted on the breadboard in the three-dimensional space. For each of the letter, 20 training instances are captured. The number of training instances can be varied (say 40 or 50). Each training instance consists of numerical values of x, y and z axis for both the accelerometer (ax, ay, az) and the gyroscope (gx, gy, gz) captured at different position in a 3-dimensional space. The Gyroscope gives the values of Angular Velocity (degrees/sec) in the three respective axes. For Accelerometer (which gives x, y, z axes acceleration due to gravity), the unit used is g ( $9.81 \text{ m/s}^2$ ). The scale of each of the sensor depends on the sensitivity settings chosen. It can be one of  $\pm 2$ , 4, 8, or 16g for the accelerometer and  $\pm 250$ , 500, 1000, or 2000 deg/sec for the gyroscope. For the proposed experiment, the default values which are already set at the components production were used. The default setting is  $\pm 2$ g for the accelerometer and  $\pm 250$  deg/sec for the gyroscope. Figure 4 shows few of the patterns used for capturing the some of the training instances for the letter 'a', 'b', 'c', 'd', 'e', 'v', 'p', 'q' and 'r'. The arrow marks indicate the hand movement direction.

2. Data Pre-processing: It means standardization of the captured data. The term 'Standardization' here refers to the shifting the distribution of each attribute to have a mean of zero and a standard deviation of one (unit variance). First, the data is converted into floating point values and then the data is scaled and the values of each axis are saved in a separate variable.
3. Features extraction (using linear interpolation): The term 'linear interpolation' is a method of curve fitting using linear polynomial to construct new data points

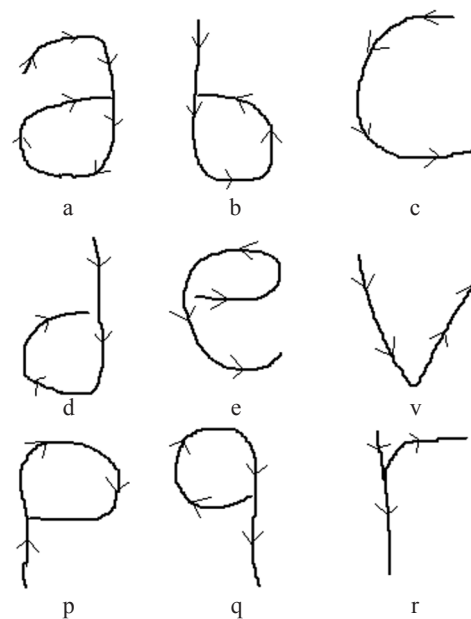


FIGURE 4. Training pattern for some of the letters of English alphabets

within a preferred range of a set of known data points. For the experiment conducted, a python library present in scipy for linear interpolation in 1-dimension is used to interpolate the data. Assuming that the obtained data signal has 70 points, then the data are fitted within the data points of size 50. Or if the obtained data points are 30, the data is again fitted within the data points of size 50. The data points of size 50 is pre-defined and can be altered. Since the size fit of 50 is taken and there are 6 axes, the total number of features for each sample data is 300. The signals generated after interpolation for each of the letter of the English alphabet pattern that we have acquired is different. Figure 5 shows the signals generated for letter 'a' for all the axes of the accelerometer and gyroscope. The first plot is for the raw data, the second plot is for the standardized data which is scaled up and the third plot is the interpolated signals for each input instance.

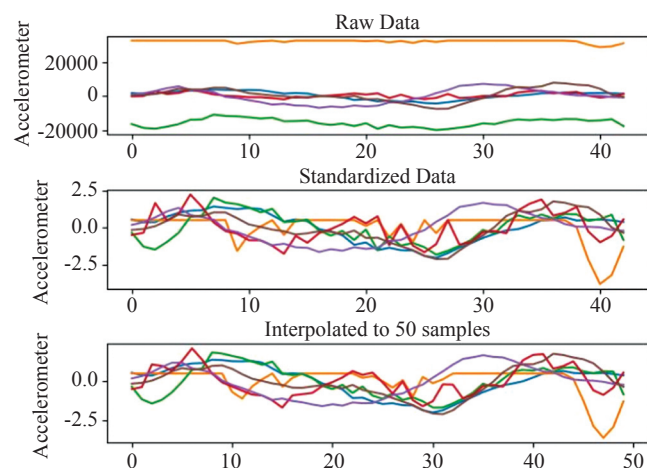


FIGURE 5. Accelerometer data signals generated for 'a'





- Ahmad, A., Feng, C. & Valaee, S. 2011. A novel accelerometer-based gesture recognition system. *IEEE Transactions on Signal Processing* 59(12): 6197-6205.
- Ajay, K., Ramesh, A., Srinivasan, L. & Vijayaraghavan, V. 2017. Low-cost static gesture recognition system using mems accelerometers. *Global Internet of Things Summit (GloTS)*, 1-6.
- Boyali, A. & Kavakli, M. 2012. A robust gesture recognition algorithm based on sparse representation, random projections and compressed sensing. *2012 7<sup>th</sup> IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 243-249.
- Gupta, H.P., Chudgar, H.S., Mukherjee, S., Dutta, T. & Sharma, K.. 2016. A continuous hand gestures recognition technique for human-machine interaction using accelerometer and gyroscope sensors. *IEEE Sensor Journal* 16(16): 6425-6432.
- Hsu, Y.L., Chu, C.L., Tsai, Y.J. & Wang, J.S. 2015. An inertial pen with dynamic time warping recognizer for handwriting and gesture recognition. *IEEE Sensors Journal* 15(1): 154-163.
- Jakub, G., Maşior, M., Zaborski, M. & Barczewska, K. 2016. Inertial motion sensing glove for sign language gesture acquisition and recognition. *IEEE Sensors Journal* 16(16): 6310-6316.
- Liu, J., Wang, Z. & Zhong, L. 2009. uWave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing* 5(6): 657-675.
- Molchanov, P., Gupta, S., Kim, K., Kautz, J. & Clara, S. 2015. Hand gesture recognition with 3d convolutional neural networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshop*, 1-7.
- Muley, P.S. & Yadav, D.M. 2014. Digital pen based on trajectory recognition algorithm. *International Journal of Engineering and Technical Research* 2(6): 22-24.
- Nagadeepa, Ch., Balaji, N. & Padmaja, V. 2016. Analysis of inertial sensor data using trajectory recognition algorithm. *International Journal on Cybernetics & Informatics* 5(4):101-108.
- Niezen, G. & Hancke, G.P. 2009. Evaluating and optimising accelerometer-based gesture recognition techniques for mobile devices. *IEEE AFRICON*, 1-6.
- Setia, A., Mittal, S., Nigam, P., Singh, S. & Gangwar, S. 2015. Hand gesture recognition based robot using accelerometer sensor. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* 4(5): 4470-4476.
- Xie, R., Sun, X., Xia, X. & Cao, J.C. 2015. Similarity matching based extensible hand gesture recognition. *IEEE Sensors Journal* 15(6): 1-1.
- Xie, R. & Cao, J. 2016. Accelerometer-based hand gesture recognition by neural network and similarity matching. *IEEE Sensors Journal* 16(11): 4537-4545.
- Xu, R., Zhou, S. & Li, W.J. 2012. MEMS accelerometer based nonspecific-user hand gesture recognition. *IEEE Sensors Journal* 12 (5): 1166-73.